Imbedding Conditions for λ-Matrices

E. Marques de Sá

Instituto de Matemática
Universidade de Coimbra
Portugal

Submitted by David H. Carlson

ABSTRACT

We say that $A(\lambda)$ is λ-imbeddable in $B(\lambda)$ whenever $B(\lambda)$ is equivalent to a λ-matrix having $A(\lambda)$ as a submatrix. In this paper we solve the problem of finding a necessary and sufficient condition for $A(\lambda)$ to be λ-imbeddable in $B(\lambda)$. The solution is given in terms of the invariant polynomials of $A(\lambda)$ and $B(\lambda)$. We also solve an analogous problem when $A(\lambda)$ and $B(\lambda)$ are required to be equivalent to regular λ-matrices. As a consequence we give a necessary and sufficient condition for the existence of a matrix B, over a field F, with prescribed similarity invariant polynomials and a prescribed principal submatrix A.

1. INTRODUCTION

All the matrices considered throughout this paper will have coefficients in the arbitrary field F, or in the polynomial ring $F[\lambda]$, in which case they will be called λ-*matrices* and denoted $A(\lambda), B(\lambda), \dots$. We use $A, B, \dots$ exclusively for matrices over F, also called constant matrices.

Imbeddings of Constant Matrices

We say that the n-square matrix A is *imbeddable* in the m-square matrix B $(m \geqslant n)$ whenever there exist two matrices $\bar{A}$ and $\bar{B}$, similar to A and B respectively, such that $\bar{A}$ is a principal submatrix of $\bar{B}$, i.e., whenever there exists an m-square nonsingular matrix U such that A is a principal submatrix of UBU^{-1}.

The sense we give here to the word "imbeddable" is similar to that of Ky Fan and Gordon Pall in [2]. There the field F is the complex number field, and U is required to be unitary. The problem that consists of determining a

 0024-3795/79/020033+18$0.175

necessary and sufficient condition of imbeddability of A in B in terms of their invariant polynomials was posed by Graciano N. de Oliveira [3]. We solve this problem in the present paper (Theorem 5.4).

Regular Imbeddings

As usual, an n-square λ-matrix $A(\lambda)$ is said to be *regular of degree* k ($k \geqslant 0$) when $A(\lambda)=\lambda^k A_0 + A_1(\lambda)$, where A_0 is nonsingular and $A_1(\lambda)$ has degree less than k. (The zero polynomial is considered to have degree less than k, although we do not define its degree.) $A(\lambda)$ is called *k-regularizable* whenever it is equivalent to a regular λ-matrix of degree k. Hence an n-square k-regularizable λ-matrix has determinant of degree nk. It turns out (Theorem 5.2) that the condition $\deg(|A(\lambda)|)=nk$ entirely characterizes the $n \times n$ k-regularizable λ-matrices.

Let $A(\lambda)$ and $B(\lambda)$ be respectively $n \times n$ and $m \times m$ ($m \geqslant n$) k-regularizable λ-matrices. $A(\lambda)$ is said to be *R-imbeddable* in $B(\lambda)$ whenever there exist two regular λ-matrices of degree k, $\bar{A}(\lambda)$ and $\bar{B}(\lambda)$, equivalent to $A(\lambda)$ and $B(\lambda)$ respectively, such that $\bar{A}(\lambda)$ is a submatrix of $\bar{B}(\lambda)$.

λ-Imbeddings

Another definition that arises naturally from the preceding ones is the following: the $n \times m$ λ-matrix $A(\lambda)$ is said to be *λ-imbeddable* in the $(n+p)\times(m+q)$ λ-matrix $B(\lambda)$ ($p,q \geqslant 0$) whenever there exist two λ-matrices $\bar{A}(\lambda)$ and $\bar{B}(\lambda)$, equivalent respectively to $A(\lambda)$ and $B(\lambda)$, such that $\bar{A}(\lambda)$ is a submatrix of $\bar{B}(\lambda)$.

In the present paper we exhibit (Theorem 3.1) a necessary and sufficient condition for $A(\lambda)$ to be λ-imbeddable in $B(\lambda)$ in terms of their invariant polynomials. We shall see also (Theorem 5.3) that R-imbeddability of k-regularizable λ-matrices is equivalent to their λ-imbeddability, and that the imbeddability of A in B is equivalent to the λ-imbeddability of the characteristic λ-matrix $\lambda I_n - A$ in $\lambda I_m - B$.

A convexity result (Lemma 4.2) seems to play an important role in passing from λ- to R-imbeddability, and the lemma may have some independent interest.

2. PRELIMINARIES

Greek letters will be used exclusively for polynomials: $\alpha, \beta, \ldots \in F[\lambda]$. We avoid the cumbersome notation $\alpha(\lambda), \beta(\lambda), \ldots$.

The Symbols $<:$, $\wedge$ *and* $\vee$.

Given two polynomials α and β, we write $\alpha <: \beta$ (or $\beta :> \alpha$) whenever there exists γ such that $\alpha\gamma = \beta$ (observe that for all α we have $0 :> \alpha$ and that $0 <: \alpha \Leftrightarrow \alpha = 0$). Moreover $\alpha \wedge \beta$ and $\alpha \vee \beta$ will stand respectively for g.c.d.(α, β) and l.c.m.(α, β) if $\alpha \neq 0$ and $\beta \neq 0$ (as usual, we take the monic determinations of the g.c.d. and the l.c.m.); we put $\alpha \vee 0 = 0$ and $\alpha \wedge 0 = \alpha'$, where α' is the monic (or null) polynomial such that $\alpha' <: \alpha <: \alpha'$. With the order relation $<:$ the set

$$\mathfrak{M} = \{\alpha : \alpha \text{ is monic or } \alpha = 0\}$$

is a complete distributive lattice, where we define $\sup(\alpha, \beta) = \alpha \vee \beta$ and $\inf(\alpha, \beta) = \alpha \wedge \beta$.

The classical identity $\alpha\beta = (\alpha \wedge \beta)(\alpha \vee \beta)$ holds in $\mathfrak{M}$.

These considerations extend to the set

$$\mathfrak{M}^Z = \{\check{\alpha} : \check{\alpha} = (\alpha_i : \alpha_i \in \mathfrak{M} \text{ for } i \in Z)\},$$

where Z is the set of integers. We define in a natural way the following order relation and operations in $\mathfrak{M}^Z$:

$$\check{\alpha} <: \check{\beta} \Leftrightarrow [\alpha_i <: \beta_i, i \in Z],$$

$$\check{\alpha} \vee \check{\beta} = \sup(\check{\alpha}, \check{\beta}) = (\alpha_i \vee \beta_i : i \in Z),$$

$$\check{\alpha} \wedge \check{\beta} = \inf(\check{\alpha}, \check{\beta}) = (\alpha_i \wedge \beta_i : i \in Z),$$

$$\check{\alpha} \cdot \check{\beta} = (\alpha_i \beta_i : i \in Z).$$

In so doing we endow $\mathfrak{M}^Z$ with the structure of a complete distributive lattice (it is in fact a product lattice), where the identity $\check{\alpha} \cdot \check{\beta} = (\check{\alpha} \wedge \check{\beta}) \cdot (\check{\alpha} \vee \check{\beta})$ holds.

Invariant Chains of λ-Matrices

We shall be concerned in the main with elements $\check{\alpha}$ of $\mathfrak{M}^Z$ such that $\alpha_i <: \alpha_{i+1}$, $i \in Z$, which will be called *chains of polynomials* or simply *chains*.

Let $A(\lambda)$ be a λ-matrix of rank r and $\alpha_1, \alpha_2, \ldots, \alpha_r$ its invariant polynomials, ordered so that $\alpha_1 <: \alpha_2 <: \ldots <: \alpha_r$. If we extend this sequence by $\alpha_i = 0$ for $i > r$ and $\alpha_i = 1$ for $i < 1$, we obtain a chain $\check{\alpha} \in \mathfrak{M}^Z$ that will be called the *invariant chain of* $A(\lambda)$. We shall also use the designation *characteristic invariant chain* of an n-square constant matrix A for the invariant chain of the characteristic λ-matrix $\lambda I_n - A$.

The Rank of $(\check{\beta})$

For an arbitrary $\check{\beta} \in \mathfrak{M}^Z$ we define $\operatorname{rank}(\check{\beta}) = \sup(i \in Z : \beta_i \neq 0) \leqslant +\infty$. Observe that $\check{\beta} <: \check{\gamma} \Rightarrow \operatorname{rank}(\check{\beta}) \geqslant \operatorname{rank}(\check{\gamma})$. If $\check{\alpha}$ is the invariant chain just defined, we have $\operatorname{rank}(\check{\alpha}) = \operatorname{rank}(A(\lambda)) = r$.

The Operator E

We introduce the shifting operators E^n, $n \in Z$, by $E^n\check{\alpha} = \check{\alpha}^*$, where $\alpha_i^* = \alpha_{i-n}$, $i \in Z$. We have then for every $n \in Z : \check{\alpha} <: \check{\beta} \Leftrightarrow E^n\check{\alpha} <: E^n\check{\beta}$ and $\operatorname{rank}(E^n\check{\alpha}) = \operatorname{rank}(\check{\alpha}) + n$.

Given a λ-matrix $A(\lambda)$ of rank r, the sequence of polynomials $\delta_k = \inf(\text{minors of order } k \text{ of } A(\lambda))$, $1 \leqslant k \leqslant r$, may be extended as follows: $\delta_k = 1$ for $k < 1$, and $\delta_k = 0$ for $k > r$. A standard result in matrix theory is that $\check{\delta} = (\delta_i : i \in Z)$ is a chain and that $\check{\delta} = \check{\alpha} \cdot E\check{\delta}$, where $\check{\alpha}$ is the invariant chain of $A(\lambda)$.

Unique Factorization in $\mathfrak{M}^Z$

Let $\pi \in \mathfrak{M}$ be an irreducible polynomial and $\alpha \in \mathfrak{M}^Z$. For each $i \in Z$ such that $\alpha_i \neq 0$, let $n_i \geqslant 0$ be the multiplicity of π in the prime factorization of α_i. We define $\check{\alpha}(\pi) = (\alpha_i(\pi) : i \in Z)$, where $\alpha_i(\pi) = \pi^{n_i}$ if $\alpha_i \neq 0$ and $\alpha_i(\pi) = 0$ if $\alpha_i = 0$. It is easily seen that every $\check{\alpha}$ may be written as a product of elements of $\mathfrak{M}^Z$:

$$\check{\alpha} = \prod_{\pi} \check{\alpha}(\pi), \tag{2.1}$$

where π runs over the set of irreducible monic polynomials. This factorization is unique up to the order of the factors. In (2.1) each coordinate α_i is represented as an infinite product of polynomials that are all zero or almost all equal to 1. A straightforward reasoning yields the following properties of $\check{\alpha}(\pi)$:

$$\check{\alpha} \text{ is a chain} \quad \Leftrightarrow \quad \text{for all } \pi, \quad \check{\alpha}(\pi) \text{ is a chain}, \tag{2.2}$$

$$\operatorname{rank}(\check{\alpha}) = \operatorname{rank}(\check{\alpha}(\pi)), \tag{2.3}$$

$$\check{\alpha} <: \check{\beta} \quad \Leftrightarrow \quad \text{for all } \pi, \quad \check{\alpha}(\pi) <: \check{\beta}(\pi), \tag{2.4}$$

$$E^n(\check{\alpha}(\pi)) = (E^n\check{\alpha})(\pi), \tag{2.5}$$

$$(\check{\alpha} \wedge \check{\beta})(\pi) = \check{\alpha}(\pi) \wedge \check{\beta}(\pi), \qquad (\check{\alpha} \vee \check{\beta})(\pi) = \check{\alpha}(\pi) \wedge \check{\beta}(\pi). \tag{2.6}$$

3. IMBEDDINGS OF POLYNOMIAL MATRICES

THEOREM 3.1. *Let $A(\lambda)$ and $B(\lambda)$ be λ-matrices of sizes $n\times m$ and $(n+p)\times(m+q)$ respectively $(p,q\geqslant 0)$, and let $\check{\alpha}$ and $\check{\beta}$ be their respective invariant chains. $A(\lambda)$ is λ-imbeddable in $B(\lambda)$ if and only if the following relations hold:*

$$E^{p+q}\check{\alpha} < :\check{\beta} < :\check{\alpha}. \tag{3.1}$$

Observe that the condition (3.1) implies that $\operatorname{rank}(\check{\alpha})\leqslant\operatorname{rank}(\check{\beta})\leqslant\operatorname{rank}(\check{\alpha})+p+q$, but $\operatorname{rank}(\check{\beta})\leqslant\min(n+p,m+q)$ is not a consequence of (3.1). The key for the proof of this theorem will be given by the following lemma.

LEMMA 3.2. *Let $K(\lambda)$ be an r-square λ-matrix of rank r and $\check{\alpha}$ its invariant chain. Let ε be a polynomial and $\check{\beta}$ an element of $\mathfrak{M}^{Z}$. Then there exists a λ-column $X(\lambda)$ of dimension r such that $\check{\beta}$ is the invariant chain of the λ-matrix*

$$\begin{bmatrix} K(\lambda) & X(\lambda) \\ 0 & \varepsilon \end{bmatrix} \tag{3.2}$$

if and only if the following conditions are satisfied:

$$E\check{\alpha} < :\check{\beta} < :\check{\alpha}, \tag{3.3}$$

$$\beta_1\beta_2\cdots\beta_{r+1} = \alpha_1\alpha_2\cdots\alpha_r\varepsilon. \tag{3.4}$$

Proof of Lemma 3.2. As $K(\lambda)$ is equivalent to $\operatorname{diag}(\alpha_1,\alpha_2,\dots,\alpha_r)$, it is clear that we can put in (3.2) $\operatorname{diag}(\alpha_1,\alpha_2,\dots,\alpha_r)$ instead of $K(\lambda)$. Our problem then consists of studying the invariant polynomials of

$$N(\lambda) = \begin{bmatrix} \alpha_1 & & & & & & \chi_1 \\ & \alpha_2 & & & & 0 & \chi_2 \\ & & \alpha_3 & & & & \chi_3 \\ & & & \ddots & & & \vdots \\ & & & & \alpha_{r-1} & & \chi_{r-1} \\ & 0 & & & & \alpha_r & \chi_r \\ 0 & 0 & 0 & \cdots & 0 & 0 & \chi_{r+1} \end{bmatrix}, \tag{3.5}$$

where the polynomials $\chi_1, \chi_2, \dots, \chi_r$ are the coordinates of $X(\lambda)$, and $\chi_{r+1} = \varepsilon$. We shall also put $\chi_0 = 1$. Let's define the following polynomials:

$$\Delta_k = \prod_{i \leqslant k} \alpha_i, \qquad k \in Z \tag{3.6}$$

$$\delta_k = \inf(\text{minors of order } k \text{ of } N(\lambda)), \qquad 1 \leqslant k \leqslant r+1 \tag{3.7}$$

$$\delta_k = 1, \quad k \leqslant 0, \quad \text{and} \quad \delta_k = 0, \quad k \geqslant r+2.$$

We split the proof into 3 steps.

1st step: we compute the δ's. Most of the minors of order k of $N(\lambda)$ are zero (at least for small k), and many of those that may be nonzero are multiples of other minors of the same order. For instance, if $1 < i_1 < i_2 < \cdots < i_{k-1} \leqslant r$, then $\chi_1 \alpha_{i_1} \alpha_{i_2} \cdots \alpha_{i_{k-1}}$ and $\chi_1 \alpha_2 \alpha_3 \cdots \alpha_k$ are minors of order k, and $\chi_1 \alpha_{i_1} \alpha_{i_2} \cdots \alpha_{i_{k-1}} : > \chi_1 \alpha_2 \alpha_3 \cdots \alpha_k$. Consequently we have to consider $r+2$ minors under the inf of (3.7), namely:

$$\begin{aligned}
\delta_k = \inf(&\alpha_1\alpha_2\alpha_3\cdots\alpha_{k-1}\alpha_k, && \chi_1\alpha_2\alpha_3\cdots\alpha_{k-1}\alpha_k,\\
&\alpha_1\chi_2\alpha_3\cdots\alpha_{k-1}\alpha_k, && \alpha_1\alpha_2\chi_3\cdots\alpha_{k-1}\alpha_k\\
&\quad\cdots, && \alpha_1\alpha_2\alpha_3\cdots\chi_{k-1}\alpha_k,\\
&\alpha_1\alpha_2\alpha_3\cdots\alpha_{k-1}\chi_k, && \alpha_1\alpha_2\alpha_3\cdots\alpha_{k-1}\chi_{k+1},\\
&\alpha_1\alpha_2\alpha_3\cdots\alpha_{k-1}\chi_{k+2}, && \quad\cdots\\
&\alpha_1\alpha_2\alpha_3\cdots\alpha_{k-1}\chi_r, && \alpha_1\alpha_2\alpha_3\cdots\alpha_{k-1}\chi_{r+1})
\end{aligned}$$

for $1 \leqslant k \leqslant r$, and $\delta_{r+1} = \Delta_r \cdot \chi_{r+1}$. This may be simplified by considering the equalities

$$\alpha_1\alpha_2\cdots\alpha_{s-1}\chi_s\alpha_{s+1}\cdots\alpha_k = \Delta_{k-1}(\alpha_k/\alpha_s)\chi_s, \qquad 1 \leqslant s \leqslant k.$$

We obtain

$$\delta_k = \Delta_{k-1}\inf\left(\frac{\alpha_k}{\alpha_0}\chi_0, \frac{\alpha_k}{\alpha_1}\chi_1, \dots, \frac{\alpha_k}{\alpha_{k-1}}\chi_{k-1}, \chi_k, \dots, \chi_{r+1}\right) \tag{3.8}$$

for $0 \leqslant k \leqslant r+1$. Now define new polynomials by

$$\theta_k = \alpha_k/\alpha_{k-1}, \qquad 1 \leqslant k \leqslant r+1,$$

$$\xi_k = \inf\left(\frac{\alpha_k}{\alpha_0}\chi_0, \frac{\alpha_k}{\alpha_1}\chi_1, \dots, \frac{\alpha_k}{\alpha_k}\chi_k\right), \qquad 0 \leqslant k \leqslant r,$$

$$\eta_k = \inf(\chi_{k+1}, \dots, \chi_r, \chi_{r+1}) \qquad 0 \leqslant k \leqslant r.$$

Because of the associativity of the inf it is easily seen that $\xi_k = \inf(\theta_k \xi_{k-1}, \chi_k)$, $1 \leqslant k \leqslant r$. Accordingly, from (3.8) we get

$$\delta_k = \Delta_{k-1} \inf(\theta_k \xi_{k-1}, \eta_{k-1}), \tag{3.9}$$

$$\delta_{k-1} = \Delta_{k-2} \inf(\xi_{k-1}, \eta_{k-1}), \qquad 1 \leqslant k \leqslant r+1. \tag{3.10}$$

2nd step: we prove the "only if" part of the lemma. Firstly we remark that, for any polynomials θ, ξ and η, we have $\inf(\theta\xi,\eta) :> \inf(\xi,\eta)$ and $\inf(\theta\xi,\eta)/\inf(\xi,\eta) <: \theta$ whenever $\xi \neq 0$ (or $\eta \neq 0$). Now if $\check{\beta}$ is the invariant chain of $N(\lambda)$, there must be

$$\beta_k = \frac{\delta_k}{\delta_{k-1}} = \alpha_{k-1} \frac{\inf(\theta_k \xi_{k-1}, \eta_{k-1})}{\inf(\xi_{k-1}, \eta_{k-1})} <: \alpha_{k-1}\theta_k = \alpha_k$$

because of (3.9) and (3.10). Therefore $\alpha_{k-1} <: \beta_k <: \alpha_k$ for $1 \leqslant k \leqslant r+1$. We have just proved (3.3) coordinatewise [for $k \leqslant 0$ and $k \geqslant r+2$ (3.3) is trivial]. As (3.4) is clear, we are done.

3rd step: we prove the "if" part of the lemma. Let $\check{\beta}$ be an element of $\mathfrak{M}^Z$ satisfying (3.3) and (3.4), i.e.,

$$\alpha_{k-1} <: \beta_k <: \alpha_k, \quad k \in Z, \quad \text{and} \quad (3.4). \tag{3.11}$$

An immediate consequence of these relations is that $\check{\beta}$ is a chain, $\beta_k = 1$ for $k \leqslant 0$, $\beta_k = 0$ for $k \geqslant r+2$, $\beta_k \neq 0$ for $k \leqslant r$, and $\beta_{r+1} = 0$ iff $\varepsilon = 0$.

We must find $\bar{\chi}_k$, $0 \leqslant k \leqslant r+1$, such that $\bar{\chi}_0 = 1$ and $\bar{\chi}_{r+1} = \varepsilon$, for which the λ-matrix (3.5) with $\chi_k = \bar{\chi}_k$ has $\check{\beta}$ as its invariant chain. We put

$$\bar{\chi}_k \doteq \prod_{i=0}^{k} \frac{\beta_i}{\alpha_{i-1}}, \qquad 0 \leqslant k \leqslant r+1. \tag{3.12}$$

First of all observe that $1 = \bar{\chi}_0 <: \bar{\chi}_1 <: \cdots <: \bar{\chi}_{r+1} = \varepsilon$ in view of (3.4) and (3.12). Next we have

$$\frac{\alpha_k}{\alpha_s}\bar{\chi}_s \prod_{s+1 \leqslant i \leqslant k} \frac{\alpha_i}{\alpha_{i-1}} \bar{\chi}_s :> \prod_{s+1 \leqslant i \leqslant k} \frac{\beta_i}{\alpha_{i-1}} \bar{\chi}_s = \bar{\chi}_k$$

for $0 \leqslant s \leqslant k \leqslant r+1$. Consequently the substitution of $\bar{\chi}_k$ for χ_k in (3.8) gives $\delta_k = \Delta_{k-1}\bar{\chi}_k$, $0 \leqslant k \leqslant r+1$, and so $\delta_k/\delta_{k-1} = \beta_k$, $1 \leqslant k \leqslant r+1$. Hence $\check{\beta}$ is the invariant chain of $N(\lambda)$, as we wished. ■

Proof of Theorem 3.1. As the theorem is trivial for $p=q=0$, we first consider the case $p+q=1$. Suppose $p=0$ and $q=1$ (for $p=1$ and $q=0$ we reverse the roles of rows and columns). Let $\operatorname{rank}(A(\lambda))=r$, and put $K(\lambda)=\operatorname{diag}(\alpha_1,\alpha_2,\dots,\alpha_r)$. The λ-matrix $A(\lambda)$ is then equivalent to

$$\begin{bmatrix} 0 & K(\lambda) \\ 0 & 0 \end{bmatrix} \tag{3.13}$$

where some of the zero blocks may not exist. If we border (3.13) with a λ-column on the right, we obtain

$$\begin{bmatrix} 0 & K(\lambda) & X(\lambda) \\ 0 & 0 & Y(\lambda) \end{bmatrix}, \tag{3.14}$$

where $X(\lambda)$ is $r\times 1$ and $Y(\lambda)=(\eta_1,\dots,\eta_{n-r})^T$. When $n>r$, $Y(\lambda)$ is equivalent to the λ-column $(\varepsilon,0,\dots,0)^T$, where $\varepsilon=\inf(\eta_1,\dots,\eta_{n-r})$. Hence (3.14) and

$$\begin{bmatrix} K(\lambda) & X(\lambda) \\ 0 & \varepsilon \end{bmatrix} \tag{3.15}$$

have the same invariant chain. [Matrices (3.14) and (3.15) are possibly nonequivalent just because they may be of different sizes.]

When $n=r$, (3.14) reduces to $[0\ K(\lambda)\ X(\lambda)]$, which has the same invariant chain as (3.15) with $\varepsilon=0$. Therefore, we are able to apply Lemma 3.2. If $\check{\beta}$ is the invariant chain of (3.14) and consequently of (3.15), then (3.3), which is just (3.1) for $p+q=1$, holds. Conversely, let's suppose that the condition (3.1), [i.e., (3.3)] is satisfied by the invariant chain $\check{\beta}$ of $B(\lambda)$. As $\beta_i :> \alpha_{i-1}$, we can define

$$\varepsilon = \beta_1(\beta_2/\alpha_1)\cdots(\beta_{r+1}/\alpha_r)$$

[Observe that in the case $n=r$ we have $\operatorname{rank}(\check{\beta})=r$ and so $\beta_{r+1}=\varepsilon=0$.] For this ε, (3.4) holds, and so Lemma 3.2 applies. The theorem is then proved for $p+q=1$.

Now the "only if" part of the theorem is evident for arbitrary p and $q\geqslant 0$ by the transitivity of $<:$. We prove the "if" part by induction on $p+q$. Suppose we are given a chain $\check{\beta}$ such that

$$E^{p+q+1}\check{\alpha} <: \check{\beta} <: \check{\alpha},$$

$$\operatorname{rank}(\check{\beta}) \leqslant \min(n+p+1, m+q) \qquad \left[\text{or} \quad \operatorname{rank}(\check{\beta}) \leqslant \min(n+p, m+q+1)\right].$$

(The last condition is of course necessary for the existence of a λ-matrix $B(\lambda)$ that is $(n+p+1)\times(m+q)$ [respectively $(n+p)\times(m+q+1)$] and whose invariant chain is $\check{\beta}$.) We must look for a chain $\check{\gamma}$ such that

$$E\check{\alpha} <: \check{\gamma} <: \check{\alpha}, \tag{3.16}$$

$$E^{p+q}\check{\gamma} <: \check{\beta} <: \check{\gamma}, \tag{3.17}$$

and

$$\operatorname{rank}(\check{\gamma}) \leqslant \min(n+1,m)$$
$$\left[\text{respectively, } \operatorname{rank}(\check{\gamma}) \leqslant \min(n,m+1)\right].$$

We just put $\check{\gamma}=\check{\alpha}\wedge E^{-p-q}\check{\beta}$. It is easily seen that for this $\check{\gamma}$, (3.16) and (3.17) are fulfilled, and that we have the following inequalities:

$$\begin{aligned}\operatorname{rank}(\check{\gamma}) &= \max\big(\operatorname{rank}(\check{\alpha}),\operatorname{rank}(\check{\beta})-p-q\big)\\ &\leqslant \max(\min(n,m),\min(n+p+1,m+q)-p-q)\\ &= \max(\min(n,m),\min(n+1-q,m-p)) \leqslant \min(n+1,m)\end{aligned}$$

[or $\operatorname{rank}(\check{\gamma}) \leqslant \min(n,m+1)$]. The theorem follows. ■

4. PATHS FOR AN IMBEDDING—SOME CONVEXITY

Let $A(\lambda)$ and $B(\lambda)$ be square λ-matrices of dimensions n and $n+p$ ($p \geqslant 0$) respectively, both of maximum rank. Suppose that $A(\lambda)$ is λ-imbeddable in $B(\lambda)$ and that $\bar{B}(\lambda)$ is a λ-matrix equivalent to $B(\lambda)$ such that $A(\lambda)=\bar{B}(\lambda)[12\ldots n|12\ldots n]$. The sequence

$$\check{\alpha} = \check{\gamma}^0,\check{\gamma}^1,\ldots,\check{\gamma}^{p-1},\check{\gamma}^p = \check{\beta}, \tag{4.1}$$

where $\check{\gamma}^i$ is the invariant chain of $\bar{B}(\lambda)[12\ldots n+i|12\ldots n+i]$, $0 \leqslant i \leqslant p$, will be called a *path from* $\check{\alpha}$ to $\check{\beta}$. By virtue of Theorem 3.1, the necessary and sufficient conditions for (4.1) to be a path from $\check{\alpha}$ to $\check{\beta}$ are

$$E^2\check{\gamma}^i <: \check{\gamma}^{i+1} <: \check{\gamma}^i \text{ and } \operatorname{rank}(\check{\gamma}^i) \leqslant n+i, \qquad \text{for} \quad 0 \leqslant i \leqslant p-1. \tag{4.2}$$

It is clear that there may exist more than one path from $\check{\alpha}$ to $\check{\beta}$, and, by Theorem 3.1, that every path satisfies $E^{2i}\check{\alpha} <: \check{\gamma}^i <: \check{\alpha}$ and $E^{2(p-i)}\check{\gamma}^i <: \check{\beta} <: \check{\gamma}^i$, or equivalently:

$$\check{\beta} \vee E^{2i}\check{\alpha} <: \check{\gamma}^i <: \check{\alpha} \wedge E^{2(i-p)}\check{\beta}, \qquad 0 \leqslant i \leqslant p. \tag{4.3}$$

The existence of paths from $\check{\alpha}$ to $\check{\beta}$ with the additional property of having maximum rank, i.e.,

$$\operatorname{rank}(\check{\gamma}^i) = n + i, \qquad 0 \leqslant i \leqslant p, \tag{4.4}$$

is provided by the following example: $\check{\mu}^i$ is the chain whose jth coordinate is the polynomial μ_j^i given by

$$\mu_j^i = \begin{cases} \beta_j \vee \alpha_{j-2i} & \text{for} \quad j \leqslant n+i, \\ 0 & \text{for} \quad j > n+i, \end{cases} \tag{4.5}$$

$i = 0, 1, \dots, p$. To see that $(\check{\mu}^i)$ is a path from $\check{\alpha}$ to $\check{\beta}$, we use the relations

$$\beta_{j-2} \vee \alpha_{j-2i-2} <: \beta_j \vee \alpha_{j-2i-2} <: \beta_j \vee \alpha_{j-2i}, \qquad \text{for all } j, \quad 0 \leqslant i \leqslant p-1,$$

which imply that

$$\mu_{j-2}^i <: \mu_j^{i+1} \text{ and } \mu_j^{i+1} <: \mu_j^i \qquad \text{for all } j, \quad 0 \leqslant i \leqslant p-1.$$

Therefore we have

$$E^2\check{\mu}^i <: \check{\mu}^{i+1} <: \check{\mu}^i \quad \text{and} \quad \operatorname{rank}(\check{\mu}^i) = n + i \tag{4.6}$$

for $0 \leqslant i \leqslant p-1$. By (4.2), $(\check{\mu}^i)$ is a path and has the required rank, as we claimed. ■

Remark. It is interesting to notice that the path $(\check{\mu}^i)$ could also be defined as

$$\check{\mu}^i = \inf(\check{\gamma}^i \in \mathfrak{M}^Z : \check{\gamma}^i \text{ is a chain satisfying (4.3) and (4.4)}), \tag{4.7}$$

$i = 0, 1, \dots, p$. In this sense $(\check{\mu}^i)$ is *the minimum path from $\check{\alpha}$ to $\check{\beta}$.*

Definition. If $\check{0} \neq \check{\gamma} \in \mathfrak{M}^Z$, let the *degree* of $\check{\gamma}$ be

$$\deg(\check{\gamma}) = \sum_i \deg(\gamma_i) \leqslant +\infty,$$

where the summation is extended to all $i \in Z$ such that $\gamma_i \neq 0$. Whenever a path (4.1) is given, we define its *degree function* as the following real-valued function whose domain is the real interval $[0,p]$:

$$D_\gamma(x) = \deg(\check{\gamma}^i) + \left[\deg(\check{\gamma}^{i+1}) - \deg(\check{\gamma}^i)\right](x-i)$$

for $x \in [i, i+1]$, $0 \leqslant i \leqslant p-1$.

It is our purpose in the present section to prove the following

PROPOSITION 4.1.

(1) *The path* $(\check{\mu}^i)$ *from* $\check{\alpha}$ *to* $\check{\beta}$ *defined by* (4.5) *or* (4.7) *has a convex degree function.*

(2) *A real function* $D(x)$, *defined on* $[0,p]$, *whose graph is a polygonal line linking vertices of integer coordinates, is the degree function of a path from* $\check{\alpha}$ *to* $\check{\beta}$ *satisfying* (4.4) *if and only if* $D(0) = \deg(\check{\alpha})$, $D(p) = \deg(\check{\beta})$ *and* $D(x) \geqslant D_\mu(x)$.

The proof of this proposition will be simplified if we prove first a lemma where functions of the type

$$F(x) = \int_e^{cx+d} \max(a(t-hx), b(t))\, dt \tag{4.8}$$

are considered for $x \in X$, X an open interval, $a(t)$ and $b(t)$ being real valued *nondecreasing* functions of the real variable t.

The functions $a(t)$ and $b(t)$ are supposed to be defined on sufficiently large open intervals so that the integrand of (4.8) is defined for each $x \in X$ and every $t \in [e, cx+d]$. (We shall eventually write $[y_1, y_2]$ instead of the usual $[y_2, y_1]$ when $y_2 < y_1$.)

LEMMA 4.2. *Let* c, d, e *and* h *be real numbers. Let* a *and* b *be nondecreasing functions as above, and suppose that* $a(t)$ *is constant for* $t < e - \inf(hX)$. *Then* $F(x)$ *defined by* (4.8) *is a convex function for* $x \in X$.

Proof. We split the proof into 3 steps.

1st step. A simple change of variables ($t' = t - e$ and $x' = -x$ if $h < 0$) shows that without loss of generality we may (as we shall do from now on) consider $e = 0$ and $h \geqslant 0$. We also remove the case $c = d = 0$, since then $F(x) \equiv 0$ is convex.

As convexity may be viewed as a local property, we have only to prove the lemma for an (arbitrary) open interval I whose closure $\bar{I}$ is contained in X, and for a and b defined on (and if necessary extended to) a suitably large compact interval K. Then a and b are $L^1 = L^1(K)$ functions.

Let $\mathcal{Q}$ be the set of pairs (a,b) of nondecreasing real-valued functions with domain K, $a(t)$ being constant for $t< -h\inf(I)$.

If (a_n) and (b_n) are sequences of L^1 functions converging (in L^1 sense) to a and b respectively, the functions

$$F_n(x) = \int_0^{cx+d} \max(a_n(t-hx), b_n(t))\,dt$$

are pointwise convergent for every $x\in I$, and the pointwise limit is a convex function whenever $F_n(x)$ is convex for all n [4, Theorem 10.8]. So we only need to prove the lemma for a subclass of $\mathcal{Q}$, $L^1\times L^1$-dense in $\mathcal{Q}$. We notice that the class $\mathcal{B}$ of the $(a,b)\in\mathcal{Q}$ such that

(1) a and b are *polygonal functions* (i.e., continuous functions whose graphs are polygonal lines with a finite number of edges),

(2) if $(t,a(t))$ and $(t',b(t'))$ are vertices of the graphs of a and b, then $a(t)\neq b(t')$,

(3) the set $\{t\in K\cap(K+y): a(t-y)=b(t)\}$ is finite for every y

is $L^1\times L^1$-dense in $\mathcal{Q}$. In fact, it is an easy exercise to prove that a nondecreasing function can be L^1-approximated by means of polygonal nondecreasing functions. Moreover, a polygonal nondecreasing function can be L^1-approximated by means of polygonal nondecreasing functions whose polygonal graphs have vertices with rational coordinates (1st type), or alternatively vertices with irrational coordinates and edges with irrational slopes (2nd type). Therefore the class of the $(a,b)\in\mathcal{Q}$ with a of the 1st type and b of the 2nd type is a subclass of $\mathcal{B}$ dense in $\mathcal{Q}$. From now on we shall be concerned with a pair $(a,b)\in\mathcal{B}$.

2nd step. We prove that $F(x)$ is everywhere differentiable. If we let $m(t,x)=\max(a(t-hx),b(t))$, an easy computation shows that

$$\frac{F(y)-F(x)}{y-x} = \int_0^{cx+d} G(t,x,y)\,dt + \frac{1}{y-x}\int_{cx+d}^{cy+d} m(t,y)\,dt \tag{4.9}$$

for $x\neq y\in I$, where $G(t,x,y)=[m(t,y)-m(t,x)]/(y-x)$.

Now define the following sets:

$$A(x) = \{t\in\,]0,cx+d[\, : a(t-hx)>b(t)\},$$

$$B(x) = \{t\in\,]0,cx+d[\, : a(t-hx)<b(t)\}. \tag{4.10}$$

As a consequence of (1), $m(t,x)$ is a continuous function whose graph (in $\mathbb{R}^3$) is a polyhedron with a finite number of faces. Therefore $\pi(t,x)$ is everywhere right and left partially differentiable with respect to either variable and has bounded derivatives. Consequently $G(t,x,y)$ is bounded [1, p. 163, Problem 3], and it is easily seen that for a fixed x

$$\lim_{y\to x} G(t,x,y) = \begin{cases} -h\cdot a'(t-hx) & \text{for} \quad t\in A(x) \quad \text{(a.e.)}, \\ 0 & \text{for} \quad t\in B(x), \end{cases} \tag{4.11}$$

(a.e. = almost everywhere), where a' is to be understood as the a.e. derivative of a. The limit (4.11) exists then for a.e. t in I because of (3). Therefore the limit of (4.9) when $y\to x$ exists [apply the Lebesgue dominated-convergence theorem to the first integral and the mean-value theorem to the second integral of (4.9)], and we have

$$\begin{aligned} F'(x) &= -h\int_{A(x)} a'(t-hx)\,dt + c\cdot m(cx+d,x) \\ &= -h\cdot V[a(t-hx):t\in A(x)] + c\cdot m(cx+d,x), \end{aligned} \tag{4.12}$$

where $V[a(t-hx):t\in A(x)]$ is the variation of $a(t-hx)$ for t running through $A(x)$. [Observe that for $cx+d\leqslant 0$ this variation is null because $a(t-hx)$ is constant for negative t.]

3rd step. We prove that $F'(x)$ is polygonal nondecreasing. Let $G_a(x)$ be the graph of $a(t-hx)$, $t\in K+hI$, and G_b the graph of $b(t)$, $t\in K$. If $h>0$, let V be the set of those $x\in I$ for which a vertex of $G_a(x)$ $[G_b]$ meets G_b $[G_a(x)]$. By (1)–(3), V is finite. If $h=0$, we let $V=\varnothing$. Therefore $I\setminus V$ is an open dense subset of I and has a finite number of connected components (c.c.). Let us fix one of those c.c. $G_a(x)\cap G_b$ has there a *fixed* and *finite* number of points. If we order them by their abcissae, the kth point, say, describes continuously and uniformly, as x varies, part of a single edge of $G_a(x)$ and of G_b. This means that if $(u(x),b(u(x)))=(u(x),a(u(x)-hx))$ are the coordinates of the point, $u(x)$ and $b(u(x))=a(u(x)-hx)$ are one-edged polygonal functions for x in the chosen c.c. Moreover, if $(v(x),b(v(x)))=(v(x),a(v(x)-hx))$ is another point of $G_a(x)\cap G_b$ such that

$$v(x)<u(x) \text{ and } a(t-hx)>b(t) \qquad \text{for} \quad v(x)<t<u(x),$$

it is easily seen that

(a) $a(u(x)-hx)$ is a nonincreasing one-edged polygonal function and
(b) $a(v(x)-hx)$ is a nondecreasing one-edged polygonal function

for x in the fixed c.c. of $I\setminus V$. The set $A(x)$ defined in (4.10) is then a union of a finite number $k(x)$ of pairwise disjoint open intervals:

$$A(x) = \bigcup_{i=1}^{k(x)} \left]v_i(x), u_i(x)\right[,$$

where $v_i(x) < u_i(x)$ for $1 \leqslant i \leqslant k(x)$, and $u_i(x) \leqslant v_{i+1}(x)$ $[u_i(x) < v_{i+1}(x)$ if $x \notin V]$ for $1 \leqslant i \leqslant k(x)-1$. We put $k(x)=0$ iff $A(x)=\varnothing$. Accordingly we have from (4.12)

$$F'(x) = h \sum_{i=1}^{k(x)} \left[a(v_i(x) - hx) - a(u_i(x) - hx) \right]$$

$$+ c \max(a(cx + d - hx), b(cx+d)). \tag{4.13}$$

For $x \in I \setminus V$ the functions $a(u_i(x) - hx)$ and $a(v_{i+1}(x) - hx)$ are certainly of the type described in (a) and (b) respectively, $1 \leqslant i \leqslant k(x)-1$. The function $a(v_1(x) - hx)$ is also a nondecreasing one-edged polygonal function in each c.c. of $I \setminus V$, even when $a(v_1(x) - hx) > b(v_1(x))$, for then $v_1(x) \leqslant 0$ and $a(v_1(x) - hx)$ is constant. To take into account the case of $u_{k(x)}$, we define the following sets:

$$\begin{aligned}
N &= \{x \in I : cx + d < 0\},\\
P &= \{x \in I : cx + d > 0,\ a(cx + d - hx) > b(cx + d)\},\\
Q &= \{x \in I : cx + d > 0,\ a(cx + d - hx) < b(cx + d)\},\\
R &= \{x \in I : cx + d > 0,\ a(cx + d - hx) = b(cx + d)\},\\
S &= \text{interior of } R.
\end{aligned} \tag{4.14}$$

It is easily seen that N, P, Q and S are pairwise disjoint open sets and that R is a finite union of compact intervals (in fact, the intersection of R with any c.c. of $I \setminus V$ either is void or is the entire c.c. or reduces to a single point). As a consequence $U = (N \cup P \cup Q \cup S) \setminus V$ is an *open dense* subset of I with a *finite* number of c.c. each entirely contained in one of the sets N, P, Q or S.

We claim that $F'(x)$ is polygonal nondecreasing in each c.c. of U. Let x be a point varying over one of those c.c. The number $k(x)$ appearing in (4.13) is constant, and by the considerations just above (4.14), we have only to prove our claim for one of the functions

$$\begin{aligned}
r_1(x) &= c \cdot b(cx + d) && \text{if } x \in Q \cup S,\\
r_2(x) &= (c - h) \cdot a((c-h)x + d) && \text{if } x \in P,\\
r_3(x) &= c \max(a(cx + d - hx), b(cx + d)) && \text{if } x \in N.
\end{aligned}$$

The $r_i(x)$, $i=1,2,3$, are clearly polygonal functions; $r_1(x)$ is nondecreasing because its right [left] derivative is $c^2(D^+b)(cx+d)\geqslant 0$ (D^+ = right derivative) if $c\geqslant 0$ [if $c\leqslant 0$]; $r_2(x)$ is nondecreasing by a similar argument. To prove our claim for the case of $r_3(x)$, just observe that $a(cx+d-hx)$ is constant if $x\in N$, and that $r_3(x)=r_1(x)$ whenever $b(cx+d)>a(cx+d-hx)$.

Finally, as $F'(x)$ is an "everywhere derivative" and is a sectionally polygonal nondecreasing function, it must be a polygonal (continuous) nondecreasing function. The lemma is proved. ∎

Proof of Proposition 4.1. (1) Let's factorize $\check{\mu}^i$ according to (2.1):

$$\check{\mu}^i = \prod_{\pi\in\mathcal{P}} \check{\mu}^i(\pi), \tag{4.15}$$

where $\mathcal{P}\subset\mathcal{M}$ is the (finite) set of all the irreducible factors of the μ_j^i, $j\leqslant n+i$, $0\leqslant i\leqslant p$. Using the properties (2.2)–(2.6) in (4.5) and (4.6), we can easily conclude that

$$\begin{aligned} &\operatorname{rank}(\check{\mu}^i(\pi))=n+i, && 0\leqslant i\leqslant p,\\ &E^2(\check{\mu}^i(\pi))<:\check{\mu}^{i+1}(\pi)<:\check{\mu}^i(\pi), && 0\leqslant i\leqslant p-1,\\ &\mu_j^i(\pi)=\beta_j(\pi)\vee\alpha_{j-2i}(\pi), \qquad j\leqslant n+i, && 0\leqslant i\leqslant p.\end{aligned}$$

for every $\pi\in\mathcal{P}$. This means that for each $\pi\in\mathcal{P}$, $(\check{\mu}^i(\pi))$ is a path from $\check{\alpha}(\pi)$ to $\check{\beta}(\pi)$ [see (4.2)], and we can even say that it is the minimum path [in the same sense of (4.7)] from $\check{\alpha}(\pi)$ to $\check{\beta}(\pi)$. By virtue of (4.15),

$$D_\mu(x) = \sum_{\pi\in\mathcal{P}} D_{\mu(\pi)}(x), \qquad x\in[0,p]$$

where $D_{\mu(\pi)}$ is the degree function of the path $(\check{\mu}^i(\pi))$. Therefore, as the sum of convex functions is convex, we only need to prove the convexity of the degree function in the case when the polynomial coordinates of $\check{\alpha}$ and $\check{\beta}$ are powers of the same irreducible polynomial. In this particular case we have

$$D_\mu(i) = \deg(\check{\mu}^i) = \sum_{j=1}^{n+i}\deg(\mu_j^i) = \sum_{j=1}^{n+i}\max\big(\deg(\alpha_{j-2i}),\deg(\beta_j)\big)$$

for $i=0,1,\dots,p$. Now apply the previous lemma to the case where $X=]0,p[$, $c=1$, $d=n$, $e=0$, $h=2$ and where $a(t)$ and $b(t)$ are defined on $\mathbb{R}^1$ by: $a(t)=b(t)=0$ for $t<0$; $a(t)=\deg(\alpha_j)$ for $j-1\leqslant t<j$, $1\leqslant j\leqslant n-1$; $a(t)=\deg(\alpha_n)$ for $t\geqslant n-1$; $b(t)=\deg(\beta_j)$ for $j-1\leqslant t<j$, $1\leqslant j\leqslant n+p-1$; $b(t)=$

$\deg(\beta_{n+p})$ for $t \geqslant n+p-1$. As the conditions of Lemma 4.2 are fulfilled, $F(x)$ defined on $[0,p]$ and given by (4.8) for $0<x<p$, $F(0)=\deg(\check{\alpha})$, and $F(p)=\deg(\check{\beta})$, is a continuous convex function. Moreover we have $F(i)=D_\mu(i)$ for $i=0,1,\ldots,p$, and so $D_\mu(x)$ is convex.

(2) We observe that the sole restrictions imposed by (4.2) on γ_{n+i}^i are: $\gamma_{n+i}^i :> \gamma_{n+i}^{i+1}$, $0 \leqslant i \leqslant p-1$, and $\gamma_{n+i}^i :> \gamma_{m+i-2}^{i-1}$, $1 \leqslant i \leqslant p$. Therefore if we multiply γ_{n+i}^i by an arbitrary $\omega_i \in \mathfrak{M}$, leaving unchanged the other coordinates of $\check{\gamma}^i$, $1 \leqslant i \leqslant p-1$, then the sequence so obtained $\check{\alpha}, \check{\gamma}_*^1, \check{\gamma}_*^2, \ldots, \check{\gamma}_*^{p-1}, \check{\beta}$, is still a path from $\check{\alpha}$ to $\check{\beta}$. The present part of the proposition follows from these observations, from the fact that $\mu_{n+i}^i \neq 0$, and from the minimum property (4.7). ∎

5. REGULAR IMBEDDINGS

LEMMA 5.1. *Let $A(\lambda)$ be an $n \times n$ λ-matrix of the form $A(\lambda)=\lambda^k I_n + A_1(\lambda)$, $k \geqslant 0$, where $A_1(\lambda)$ is a λ-matrix of degree $\leqslant k-1$. Let $\check{\alpha}$ be the invariant chain of $A(\lambda)$ and $\check{\beta}$ a chain such that $E^{2p}\check{\alpha} <: \check{\beta} <: \check{\alpha}$, rank$(\check{\beta})=n+p$ and $\deg(\check{\beta})=(n+p)k$, $p \geqslant 0$. Then there exists an $n+p$-square λ-matrix $B(\lambda)$ of the form $B(\lambda)=\lambda^k I_{n+p}+B_1(\lambda)$, where $B_1(\lambda)$ has degree $\leqslant k-1$, having $\check{\beta}$ as invariant chain and $A(\lambda)$ as a (principal) submatrix.*

Proof. For $p=0$ there's nothing to prove. We firstly consider the case $p=1$. By Theorem 3.1 there exists a λ-column $X(\lambda)$ and a λ-row $Y(\lambda)$ and a polynomial ρ such that

$$\bar{B}(\lambda) = \begin{bmatrix} A(\lambda) & X(\lambda) \\ Y(\lambda) & \rho \end{bmatrix}$$

has $\check{\beta}$ as invariant chain. We can also suppose that $\bar{B}(\lambda)$ is "normalized" so that $|\bar{B}(\lambda)|$ is a monic polynomial. As right and left divisions with remainder by $A(\lambda)$ are always possible [and unique because $A(\lambda)$ is regular], let $X(\lambda)=A(\lambda)C(\lambda)+R(\lambda)$ and $Y(\lambda)=L(\lambda)A(\lambda)+S(\lambda)$, where $C(\lambda)$ and $R(\lambda)$ are λ-columns, $L(\lambda)$ and $S(\lambda)$ are λ-rows, $R(\lambda)=(\rho_1,\rho_2,\ldots,\rho_n)^T$ and $S(\lambda)=(\sigma_1,\sigma_2,\ldots,\sigma_n)$ having degrees $\leqslant k-1$. The matrix

$$B(\lambda) = \begin{bmatrix} A(\lambda) & R(\lambda) \\ S(\lambda) & \omega \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -L(\lambda) & 1 \end{bmatrix} \bar{B}(\lambda) \begin{bmatrix} I_n & -C(\lambda) \\ 0 & 1 \end{bmatrix}$$

is equivalent to $\bar{B}(\lambda)$, and we have $|B(\lambda)|=\beta_1\beta_2\ldots\beta_{n+1}$. Hence $\deg(|B(\lambda)|)=\deg(\check{\beta})=(n+1)k$. If we develop $|B(\lambda)|$ successively by the $(n+1)$th column and row, we obtain

$$|B(\lambda)| = \omega|A(\lambda)| - \sum_{i,j=1}^{n} (-1)^{i+j}\sigma_i\rho_j|A(\lambda)(j|i)|.$$

As $\deg(\sigma_i\rho_j|A(\lambda)(j|i)|) \leqslant 2(k-1)+(n-1)k<(n+1)k$, we must have $\deg(\omega|A(\lambda)|)=(n+1)k$. Therefore ω is a monic polynomial of degree k, and so $B(\lambda)$ satisfies the requirements of the lemma.

For a general p, we observe that as $\deg(\check{\alpha})=nk$, $\deg(\check{\beta})=(n+p)k$, $\operatorname{rank}(\check{\alpha})=n$, $\operatorname{rank}(\check{\beta})=n+p$, and as (3.1) holds, then by Theorem 3.1, the definition of $(\check{\mu}^i)$ and Proposition 4.1, statements (1) and (2), there exists a path $(\check{\gamma}^i)$ from $\check{\alpha}$ to $\check{\beta}$ satisfying (4.4), whose degree function is linear, i.e., $D_\gamma(i)=\deg(\check{\gamma}^i)=(n+i)k$. The case $p=1$ just proved applies to the sequence of the $\check{\gamma}^i$'s, and the general case follows by an easy induction. ∎

THEOREM 5.2. *An m-square λ-matrix $B(\lambda)$ is k-regularizable* ($k\geqslant 0$) *if and only if* $\deg(|B(\lambda)|)=mk$.

Proof. It is the "if" part that we must prove. If $\check{\beta}$ is the invariant chain of $B(\lambda)$, β_1 has at most degree k. Let $\alpha\in\mathfrak{M}$ have degree k and $\alpha:>\beta_1$. Put $A(\lambda)=[\alpha]$, and apply the previous lemma to $A(\lambda)$ and $B(\lambda)$ with $n=1$ and $p=m-1$. ∎

As easy consequences of Lemma 5.1 we can state the following two theorems.

THEOREM 5.3. *Let $A(\lambda)$ and $B(\lambda)$ be $n\times n$ and $(n+p)\times(n+p)$ k-regularizable λ-matrices whose invariant chains are $\check{\alpha}$ and $\check{\beta}$, respectively. $A(\lambda)$ is R-imbeddable in $B(\lambda)$ if and only if $A(\lambda)$ is λ-imbeddable in $B(\lambda)$, i.e., if and only if $E^{2p}\check{\alpha}<:\check{\beta}<:\check{\alpha}$.*

THEOREM 5.4. *Let A and B be $n\times n$ and $(n+p)\times(n+p)$ constant matrices whose characteristic invariant chains are $\check{\alpha}$ and $\check{\beta}$, respectively. A is imbeddable in B if and only if $\lambda I_n - A$ is (R-) λ-imbeddable in $\lambda I_{n+p}-B$, i.e., if and only if $E^{2p}\check{\alpha}<:\check{\beta}<:\check{\alpha}$.*

6. REMARK ON THOMPSON'S "INTERLACING INEQUALITIES FOR INVARIANT FACTORS"

Several months after this article was submitted, we learned that Professor R. C. Thompson [5] had also solved the same problems that are solved here. We discuss briefly the similarities and differences between the two articles.

First, where we characterize the invariant polynomials of the submatrices of a given λ-matrix, Thompson deals with the same question for matrices over arbitrary principal ideal domains; all our work here can be easily extended to this setting.

Second, in the proof of our Lemma 3.2, where we deal with divisibility directly, he localizes the problem, transforming it into one involving systems of integer inequalities.

Third, where we discuss imbeddings of a regular λ-matrix, Thompson deals with a characteristic matrix $\lambda I - A$. The degree restriction involved here is the source of many difficulties. His Lemma 3 may be regarded as a discrete version of our somewhat stronger Lemma 4.2; the consideration of this continuous version allows us to use the differential characterization of convexity. We think that our methods give some additional insight into the role of convexity in this problem.

REFERENCES

1 J. Dieudonné, *Éléments d'Analyse*, Tome 1, Gauthier-Villars, Paris, 1969.
2 Ky Fan and G. Pall, Imbedding conditions for hermitian and normal matrices, *Canad. J. Math.* 9:298–304 (1957).
3 G. N. de Oliveira, Matrices with prescribed characteristic polynomial and a prescribed submatrix III, *Monatsh. Math.* 75:441–446 (1971).
4 R. Tyrrell Rockafellar, *Convex Analysis*, Princeton U. P., 1970.
5 R. C. Thompson, Interlacing inequalities for invariant factors, *Linear Algebra and Appl.*, to appear.

Received 24 May 1977; revised 27 December 1977